418 Rec'd PCT/PTO 13 AUG 1999

FORM PTO-1390 (Modified) (REV 10-95) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER
A30919-PCT USA

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR
To be assigned 09/367433

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US98/02668	**13 February 1998**	**14 February 1997**

TITLE OF INVENTION
OBJECT-BASED AUDIO-VISUAL TERMINAL AND BITSTREAM STRUCTURE

APPLICANT(S) FOR DO/EO/US
ELEFTHERIADIS, Alexandros and KALVA, Hari

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   - a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
   - b. ☐ has been transmitted by the International Bureau.
   - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☒ A copy of the International Search Report (PCT/ISA/210).
8. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - b. ☐ have been transmitted by the International Bureau.
   - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - d. ☒ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
10. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 13 to 18 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☐ A **FIRST** preliminary amendment.
    A **SECOND** or **SUBSEQUENT** preliminary amendment.
16. ☐ A substitute specification.
17. ☐ A change of power of attorney and/or address letter.
18. ☒ Certificate of Mailing by Express Mail
19. ☒ Other items or information:

Form PCT/RO/101; Forms PCT/IB/301/304/308/332; Forms PCT/ISA/210/220; Forms PCT/IPEA/401/402/408; and a check in the amount of $335.00.

Express Mail Label No.: EJ339575675US

514 Rec'd PCT/PTO 13 AUG 1999

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR) To be assigned 09/737433	INTERNATIONAL APPLICATION NO. PCT/US98/02668	ATTORNEY'S DOCKET NUMBER A30919-PCT USA

20. The following fees are submitted:.

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

		CALCULATIONS	PTO USE ONLY
☐ Search Report has been prepared by the EPO or JPO	$840.00		
☒ International preliminary examination fee paid to USPTO (37 CFR 1.482)	$670.00		
☐ No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2))	$760.00		
☐ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2) paid to USPTO	$970.00		
☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4)	$96.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =		**$670.00**	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).		**$0.00**	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	10 - 20 =	0	x $18.00	**$0.00**	
Independent claims	- 3 =	0	x $78.00	**$0.00**	
Multiple Dependent Claims **(check if applicable).**			☐	**$0.00**	
TOTAL OF ABOVE CALCULATIONS =				**$670.00**	
Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) **(check if applicable).**			☒	**$335.00**	
SUBTOTAL =				**$335.00**	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). +				**$0.00**	
TOTAL NATIONAL FEE =				**$335.00**	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) **(check if applicable).**			☐	**$0.00**	
TOTAL FEES ENCLOSED =				**$335.00**	
				Amount to be: refunded	$
				charged	$

☒ A check in the amount of **$335.00** to cover the above fees is enclosed.

☐ Please charge my Deposit Account No. in the amount of to cover the above fees.
A duplicate copy of this sheet is enclosed.

☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **02-4377** A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Henry Tang, Esq,
Baker & Botts, LLP
30 Rockefeller Plaza
New York, NY 10112-0228

Peter A. Businger
SIGNATURE

Peter A. Businger
NAME

27,862
REGISTRATION NUMBER

13 August 1999
DATE

OIPE JAN 13 2000 PATENT & TRADEMARK OFFICE

Attorney's Docket No. A30919 PCT USA
Baker & Botts, L.L.P.

Applicant or Patentee: Eleftheriadis et al.
Serial or Patent No.: 09/367,433 Filed or Issued: August 13, 1999
For: OBJECT-BASED AUDIO-VISUAL TERMINAL AND BITSTREAM STRUCTURE

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) and 1.27(d)) - NONPROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

NAME OF ORGANIZATION The Trustees of Columbia University in the City of New York

ADDRESS OF ORGANIZATION 116th Street & Broadway, New York, NY 10027

TYPE OF ORGANIZATION

[X] UNIVERSITY OR OTHER INSTITUTION OF HIGHER EDUCATION
[ ] TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE]26 USC 501(a) and 501(c)(3))
[ ] NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA
(NAME OF STATE __)
(CITATION OF STATUTE __)
[ ] WOULD QUALIFY AS TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE (26 USC 501(a) and 501(c)(3)) IF LOCATED IN THE UNITED STATES OF AMERICA
[ ] WOULD QUALIFY AS NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA IF LOCATED IN THE UNITED STATES OF AMERICA
(NAME OF STATE __)
(CITATION OF STATUTE __)

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 CFR 1.9(e) for purposes of paying reduced fees under Section 41(a) and (b) of Title 35, United States Code with regard to the invention entitled OBJECT-BASED AUDIO-VISUAL TERMINAL AND BITSTREAM STRUCTURE by inventor(s) Alexandros Eleftheriadis and Hari Kalva
described in

[ ] the specification filed herewith
[X] Application Serial No. 09/367,433 , filed August 13, 1999 .
[ ] Patent No. __, issued __.

I hereby declare that the rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9(c) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e). *Note: Separate verified statements are required from each named person, concern or organization having rights to the invention

averring to their status as small entities. (37 CFR 1.27)

NAME ______________________________

ADDRESS ______________________________

[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

NAME ______________________________

ADDRESS ______________________________

[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING Jack M. Granowitz

TITLE IN ORGANIZATION Executive Director, Columbia Innovation Enterprise

ADDRESS OF PERSON SIGNING Columbia University, Engineering Terrace - 363, New York, NY 10027-6699

SIGNATURE Jack M. Granowitz Rd. DATE 12/8/99

WO 98/36559

09/367433
PCT/US98/02668
10 Rec'd PCT/PTO 13 AUG 1999

OBJECT-BASED AUDIO-VISUAL TERMINAL AND BITSTREAM STRUCTURE

Technical Field

This invention relates to the representation, transmission, processing and display of video and audio-visual information, more particularly of object-based information.

Background of the Invention

Image and video compression techniques have been developed which, unlike traditional waveform coding, attempt to capture high-level structure of visual content. Such structure is described in terms of constituent "objects" which have immediate visual relevancy, representing familiar physical objects, e.g. a ball, a table, a person, a tune or a spoken phrase. Objects are independently encoded using a compression technique that gives best quality for each object. The compressed objects are sent to a terminal along with composition information which tells the terminal where to position the objects in a scene. The terminal decodes the objects and positions them in the scene as specified by the composition information. In addition to yielding coding gains, object-based representations are beneficial with respect to modularity, reuse of content, ease of manipulation, ease of interaction with individual image components, and integration of natural, camera-captured content with synthetic, computer-generated content.

Summary of the Invention

In a preferred architecture, structure or format for information to be processed at an object-based video or audio-visual (AV) terminal, an object-oriented bitstream includes objects, composition information, and

scene demarcation information. The bitstream structure allows on-line editing, e.g. cut and paste, insertion/deletion, grouping, and special effects.

In the preferred architecture, in the interest of ease of editing, AV objects and their composition information are transmitted or accessed on separate logical channels (LCs). The architecture also makes use of "object persistence", taking advantage of some objects having a lifetime in the decoder beyond their initial presentation time, until a selected expiration time.

Brief Description of the Drawing

Fig. 1 is a functional schematic of an exemplary object-based audio-visual terminal.

Fig. 2a is a schematic of an exemplary object-based audio-visual composition packet.

Fig. 2b is a schematic of an exemplary object-based audio-visual data packet.

Fig. 2c is a schematic of an exemplary compound composition packet.

Fig. 3 is a schematic of exemplary node and scene description information using composition.

Fig. 4 is a schematic of exemplary stream-node association information.

Fig. 5 is a schematic of exemplary node/graph update information using a scene.

Fig. 6 is a schematic of an exemplary audio-visual terminal design.

Fig. 7 is a schematic of an exemplary audio-visual system controller in the terminal according to Fig. 6.

Fig. 8 is a schematic of exemplary information flow in the controller according to Fig. 7.

Detailed Description

An audio-visual (AV) terminal is a systems component which is instrumental in forming, presenting or displaying audio-visual content. This includes (but is not limited to) end-user terminals with a monitor screen and loudspeakers, as well server and mainframe computer facilities in which audio-visual information is processed. In an AV terminal, desired functionality can be hardware-, firmware- or software-implemented. Information to be processed may be furnished to the terminal from a remote information source via a telecommunications channel, or it may be retrieved from a local archive, for example. An object-oriented audio-visual terminal more specifically receives information in the form of individual objects, to be combined into scenes according to composition information supplied to the terminal.

Fig. 1 illustrates such a terminal, including a de-multiplexer (DMUX) 1 connected via a logical channel LC0 to a system controller or "executive" 2 and via logical channels LC1 through LCn to a buffer 3. The executive 2 and the buffer 3 are connected to decoders 4 which in turn are connected to a composer unit 5. Also, the executive 2 is connected to the composer unit 5 directly, and has an external input for user interaction, for example.

In the preferred AV architecture, the AV objects and their composition information are transmitted or accessed on separate logical channels. The DMUX receives the Mux2 layer from the lower layers and de-multiplexes it into logical channels. LC0 carries composition information which is passed on to the executive. The AV objects received on other logical channels are stored in the buffer to be acted upon by the

decoders. The executive receives the composition information, which includes the decoding and presentation time stamps, and instructs the decoders and composer accordingly.

The system handles object composition packets (OCP) and object data packets (ODP). A composition packet contains an object's ID, time stamps and the "composition parameters" for rendering the object. An object data packet contains an object ID, an expiration time stamp in case of persistent objects, and object data.

Preferably, any external input such as user interaction is converted to OCP and/or ODP before it is presented to the executive. There is no need for headers in a bitstream delivered over a network. However, headers are required when storing an MPEG4 presentation in a file.

Figs. 2a and 2b illustrate the structure of composition and data packets in further detail. Relevant features are as follows:

Object ID is composed of object type and object number. The default length of the Object ID is 2 bytes, including ten bits for the object number and 6 for the object type (e.g. text, graphics, MPEG2 VOP, compound object). An extensible code is used to accommodate more than 1023 objects or more than 31 object types. The following convention will be adhered to: a value of 0b111111 in the first six bits of the Object ID corresponds to 31 plus the value of the byte immediately following the ObjectID; a value of 0b11.1111.1111 in the least significant 10 bits of the Object ID corresponds to 1023 plus the value of the two bytes immediately following the Object ID (without counting the object type extension bytes, if present). The following object types

are defined:

<u>Composition Objects</u> (16-bit object IDs)

0X0000	scene configuration object
0X0001	node hierarchy specification
0X0002	stream-node association
0X0003	node/scene update
0X0004	compound object

<u>Object Data</u> (object type, 6 most significant bits)

0b00.0010	text
0b00.0011	MPEG2 VOP (rectangular VOP)

<u>Persistent Objects</u> (PO) are objects that should be saved at the decoder for use at a later time. An expiration time stamp (ETS) gives the life of a PO in milliseconds. A PO is not available to the decoder after ETS runs out. When a PO is to be used at a later time in a scene, only the corresponding composition information needs to be sent to the AV terminal.

<u>Decoding Time Stamp</u> (DTS) indicates the time an object (access unit) should be decoded by the decoder.

<u>Presentation Time Stamp</u> (PTS) indicates the time an object (access unit) should be presented by the decoder.

<u>Lifetime Time Stamp</u> (LTS) gives the duration (in milliseconds) an object should be displayed in a scene. LTS is implicit in some cases, e.g. in a video sequence where a frame is displayed for 1/frame-rate or until the next frame is available, whichever is larger. An explicit LTS is used when displaying graphics and text. An AV object should be decoded only once for use during its life time.

<u>Expiration Time Stamp</u> (ETS) is specified to support the notion of object persistence. An object, after it is presented, is saved at the decoder (cache) until a time given by ETS. Such an object can be used multiple times before ETS runs out. A PO with an expired ETS is no

longer available to the decoder.

Object Time Base (OTB) defines the notion of time of a given AV object encoder. Different objects may belong to different time bases. The AV terminal adapts these time bases to the local one, as specified in the MSDL VM.

Object Clock Reference (OCR) can be used if necessary to convey the speed of the OTB to the decoder. By this mechanism, OTBs can be recovered/adapted at the AV terminal.

Composition Parameters are used to compose a scene (place an object in a scene). These include displacement from the upper left comer of the presentation frame, rotation angles, zooming factors, etc.

Priority indicates the priority of an object for transmission, decoding, and display. MPEG4 supports 32 levels of priority. Lower numbers indicate higher priorities.

Persistence Indicator (PI) indicates whether an object is persistent.

Continuation Indicator (CI) indicates the end of an object in the current packet (or continuation).

Object Grouping facilitates operations to be applied to a set of objects with a single operation. Such a feature can be used to minimize the amount of composition information sent, as well as to support hierarchical scene composition based on independent sub-scenes. The composer manipulates the component objects as a group. The structure of a compound composition packet (CCP) is shown in Fig. 2c.

Bitstream Structure includes object composition packets for describing the composition and controlling the presentation of those packets, and object data packets that contain the data for the objects. A scene is composed by a set of composition packets. The

bitstream supports representation of scenes as a hierarchy by using compound composition objects (CCP), also known as node hierarchy. A CCP allows combining composition objects to create complex audio-visual scenes.

<u>Object-Data Packet</u>:

ObjectID - min (default) 10 bits

CI and PI could be combined:

00 - begin non-persistent
01 - begin persistent
10 - continuation
11 - end of object

Priority: 5 bits, present only if CI/PI is 0b00 or 0b01

ETS: 30 bits, present if CI/Pl is 0b01

For prediction-based video coding, VOP_type is indicated by two bits (00 (I), 01 (P), 10 (B), 11 (PB)), facilitating editing.

```
Object_data_packet{
    ObjectID                    16 bits + any extensions;
    CIPI                        2 bits
    if  (CIPI <= 1){
    Priority                    5 bits
    if (object type is MPEG VOP)
                      (any prediction based compression)
         VOP_type                     2 bits
    }
    if (CIPI == 1)
    ETS                         28  bits
    ObjectData
}
```

<u>Object Composition Packet</u>

```
Object_composition_packet{
    ObjectID                    16 bits + any extensions
    OCR_Flag                    1 bit
```

```
        Display_Timers_Flag        1 bit
        DTS                        30  bits
        if (OCR_Flag)
              OCR                  30  bits
        if (Display_Timers_Flag){
              PTS                  30  bits
              LTS                  28  bits
        }
        Composition_parameters;
}
```

Composition Parameters are defined in section 2 of MSDL Verification Model, MPEG N1483, Systems Working Draft V2.0, the disclosure of which is incorporated herein by reference.

```
Composition_parameters(
      visibility                           1 bit
            composition_order              5 bits
            number_of_motion_sets          2 bits
            x_delta_0                      12 bits
            y_delta_0                      12 bits
            for  (i = 1;  i <= number_of_motion_sets; i++){
                  x_delta_i                12 bits
                  y_delta_i                12 bits
            }
}
```

Compound Composition Packet

```
Compound_composition_packet{
      ObjectID                             16 bits
      PTS                                  30 bits
      LTS                                  28 bits
      Composition_parameters
      ObjectCount                          8 bits
      for   (i = 0; i < ObjectCount; i++){
                  Object_composition_packet;
```

```
    }
}
```

Scene Configuration Packet (SCP) is used to change reference scene width, height, to flush the buffer, and other configuration functions. The object type for SCPs is 0b00.0000. This allows for 1024 different configuration packets. The object number 0b00.0000.0000 (object ID 0X0000) is defined for use with flushing the terminal buffers.

Composition Control for Buffer Management (Object ID 0x0000)

AV terminal buffers are flushed using Flush_Cache and Scene_Update flags. When using hierarchical scene structure, the current scene graph is flushed and the terminal loads the new scene from the bitstream. Use of flags allows for saving the current scene structure instead of flushing it. These flags are used to update the reference scene width and height whenever a new scene begins. If the Flush_Cache_Flag is set, the cache is flushed, removing the objects (if any). If Scene_Update_Flag is set, there are two possibilities: (i) Flush_Cache-Flag is set, implying that the objects in the cache will no longer be used; (ii) Flush_Cache_Flag is not set, the new scene being introduced (an editing action on the bitstream) splices the current scene and the objects in the scene will be used after the end of the new scene. The ETS of the objects, if any, will be frozen for the duration of the new scene introduced. The beginning of the next scene is indicated by another scene configuration packet.

```
Scene_configuration_packet{
    ObjectID                        16 bits (OX0000)
        Flush_Cache_Flag            1 bit
```

```
        Scene_Update_Flag           1 bit
        if (Scene_Update_Flag){
            ref_scene_width         12 bits
            ref_scene_height        12 bits
        }
}
```

Composition Control for Scene Description (Object ID 0x0001)

A hierarchy of nodes is defined, describing a scene. The scene configuration packets can also be used to define a scene hierarchy that allows for a description of scenes as a hierarchy of AV objects. Each node in such a graph is a grouping of nodes that groups the leaves and/or other nodes of the graph into a compound AV object. Each node (leaf) has a unique ID followed by its parameters as shown in Fig. 3.

Composition Control for Stream-Node Mapping (Object ID 0x0002)

As illustrated by Fig. 4, table entries associate the elementary object streams in the logical channels to the nodes in a hierarchical scene. The stream IDs are unique, but not the node IDs. This implies that more than one stream can be associated with the same node.

Composition Control for Scene Updates (Object ID 0x0003)

Fig. 5 illustrates updating of the nodes in the scene hierarchy, by modifying the specific parameters of the node. The graph itself can be updated by adding/deleting the nodes in the graph. The update type in the packet indicates the type of update to be performed on the graph.

Architectural Embodiment

The embodiment described below includes an

object-based AV bitstream and a terminal architecture. The bitstream design specifies, in a binary format, how AV objects are represented and how they are to be composed. The AV terminal structure specifies how to decode and display the objects in the binary bitstream.

AV Terminal Architecture

Further to Fig. 1 and with specific reference to Fig. 6, the input to the de-multiplexer 1 is an object-based bitstream such as an MPEG-4 bitstream, consisting of AV objects and their composition information multiplexed into logical channels (LC). The composition of objects in a scene can be specified as a collection of objects with independent composition specification, or as a hierarchical scene graph. The composition and control information is included in LC0. The control information includes control commands for updating scene graphs, reset decoder buffers etc. Logical channels 1 and above contain object date. The system includes a controller (or "executive") 2 which controls the operation of the AV terminal.

The terminal further includes input buffers 3, AV object decoders 4, buffers 4' for decoded data, a composer 5, a display 6, and an object cache 7. The input bitstream may be read from a network connection or from a local storage device such as a DVD, CD-ROM or computer hard disk. LC0 containing the composition information is fed to the controller. The DMUX stores the objects in LC1 and above at the location in the buffer specified by the controller. In the case of network delivery, the encoder and the stream server cooperate to ensure that the input object buffers neither overflow nor underflow. The encoded data objects are stored in the input data buffers until read by the decoders at their decoding time, typically given by an

associated decoding timestamp. Before writing a data object to the buffer, the DMUX removes the timestamps and other headers from the object data packet and passes them to the controller for signaling of the appropriate decoders and input buffers. The decoders, when signaled by the controller, decode the data in the input buffers and store them in the decoder output buffers. The AV terminal also handles external input such as user interaction.

In the object cache 7, objects are stored for use beyond their initial presentation time. Such objects remain in the cache even if the associated node is deleted from the scene graph, but are removed only upon the expiration of an associated time interval called the expiration time stamp. This feature can be used in presentations where an object is used repeatedly over a session. The composition associated with such objects can be updated with appropriate update messages. For example, the logo of the broadcasting station can be downloaded at the beginning of the presentation and the same copy can be used for repeated display throughout a session. Subsequent composition updates can change the position of the logo on the display. Objects that are reused beyond their first presentation time may be called persistent objects.

System Controller(SC)

The system controller controls decoding and playback of bitstreams on the AV terminal. At startup, from user interaction or by looking for a session at default network address, the SC first initializes the DMUX to read from a local storage device or a network port. The control logic is loaded into the program RAM at the time of initialization. The instruction decoder reads the instructions from the program and executes

them. Execution may involve reading the data from the input buffers (composition or external data), initializing the object timers, loading or updating the object tables to the data RAM, loading object timers, or control signaling.

Fig. 7 shows the system controller in further detail. The DMUX reads the input bitstream and feeds the composition data on LC0 to the controller. The composition data begins with the description of the first scene in the AV presentation. This scene can be described as a hierarchical collection of objects using compound composition packets, or as a collection of independent object composition packets. A table that associates the elementary streams with the nodes in the scene description immediately follows the scene description. The controller loads the object IDs (stream IDs) into object list and render list which are maintained in the data RAM. The render list contains the list of objects that are to be rendered on the display device. An object that is disenabled by user interaction is removed from the render list. A node delete command that is sent via a composition control packet causes the deletion of the corresponding object IDs from the object list. The node hierarchy is also maintained in the data RAM and updated whenever a composition update is received.

The composition decoder reads data from the composition and external data buffer and converts them into a format understood by the instruction decoder. The external input includes user interaction to select objects, disenable and enable objects and certain predefined operations on the objects. During the execution of the program, two lists are formed in the data RAM. The object list, containing a list of objects

(object IDs) currently handled by the decoders and a render list, containing the list of active objects in the scene. These lists are updated dynamically as the composition information is received. For example, if a user chooses to hide an object by passing a command via the external input, the object is removed from the render list until specified by the user. This is also how external input is handled by the system. Whenever there is some external interaction, the composition decoder reads the external data buffer and performs the requested operation.

The SC also maintains timing for each AV object to signal the decoders and decoder buffers of decoding and presentation time. The timing information for the AV objects is specified in terms of its time-base. The terminal uses the system clock to convert an object's time base into system time. For objects that do not need decoding, only presentation timers are necessary. These timers are loaded with the decoding and presentation timestamps for that AV object. The controller obtains the timestamps from the DMUX for each object. When a decoding timer for an object runs out, the appropriate decoder is signaled to read data from the input buffers and to start the decoding process. When a presentation timer runs out, the decoded data for that object is transferred to the frame buffer for display. A dual buffer approach could be used to allow writing to a frame buffer while the contents of the second buffer are displayed on the monitor. The instruction decoder can also reset the DMUX or input buffers by signaling a reset, which initializes them to the default state.

Information Flow in the Controller

Fig. 8 shows the flow of information in the controller. To keep the figure simple, the operations

performed by the instruction decoder are shown in groups. The three groups respectively concern object property modifications, object timing, and signaling.

Object Property Modifications

These operations manipulate the object IDs, also called elementary stream IDs. When a scene is initially loaded, a scene graph is formed with the object IDs of the objects in the scene. The controller also forms and maintains a list of objects in the scene (object list) and active objects in the object from the render list. Other operations set and update object properties such as composition parameters when the terminal receives a composition packet.

Object Timing

This group of operations deals with managing object timers for synchronization, presentation, and decoding. An object's timestamp specified in terms of its object time base is converted into system time and the presentation and decoding time of that object are set. These operations also set and reset expiration timestamps for persistent objects.

Signaling

Signaling operations control the over-all operation of the terminal. Various components of the terminal are set, reset and operated by controller signaling. The controller checks the decoding and presentation times of the objects in the render list and signals the decoders and presentation frame buffers accordingly. It also initializes the DEMUX for reading from a network or a local storage device. At the instigation of the controller, decoders read the data from the input buffers and pass the decoded data to decoder output buffers. The decoded data is moved to the presentation device when signaled by the controller.

WE CLAIM:

1. A method for displaying object-based audio-visual/video data, comprising:

(a) receiving, over time, a plurality of audio-visual/video objects and composition information for the objects;

(b) storing in a cache memory at least one of the objects;

(c) composing scenes from said objects including the one of the objects stored in the cache memory; and

(d) displaying the composed scenes.

2. The method of claim 1, further comprising, in addition to storing the one of the objects, storing expiration time data for the one of the objects.

3. The method of claim 1, with at least one of the objects being received from a network connection.

4. The method of claim 1, with at least one of the objects being received from local memory.

5. The method of claim 1, with at least one of the objects being received from local memory and at least one other of the objects being received from a network connection, and with the composed scenes comprising the one and the other of the objects.

6. The method of claim 1, further comprising responding to interactive user input.

7. The method of claim 6, wherein responding

comprises at least one of selecting, enabling and disenabling one of the objects.

8. Apparatus for displaying object-based audio-visual/video data, comprising:

(a) a controller circuit for controlling acquisition over time of a plurality of audio-visual/video objects and composition information for the objects;

(b) a cache memory for storing at least one of the objects;

(c) a composer circuit, coupled to the cache memory, for composing scenes from said video objects including the one of the objects stored in the cache memory; and

(d) a display for the composed scene.

9. Apparatus for displaying object-based audio-visual/video data, comprising a processor which is instructed for:

(a) controlling acquisition over time of a plurality of audio-visual/video objects and composition information for the objects;

(b) storing in a cache memory at least one of the objects;

(c) composing scenes from said video objects including the one of the objects stored in the cache memory; and

(d) displaying the composed scenes.

10. Apparatus for displaying object-based audio-visual/video data, comprising:

(a) means for controlling acquisition over time of a plurality of audio-visual/video objects and

composition information for the objects;

(b) means for for storing in a cache memory at least one of the objects;

(c) means, coupled to the cache memory, for composing scenes from said video objects including the one of the objects stored in the cache memory; and

(d) means for displaying the scene.

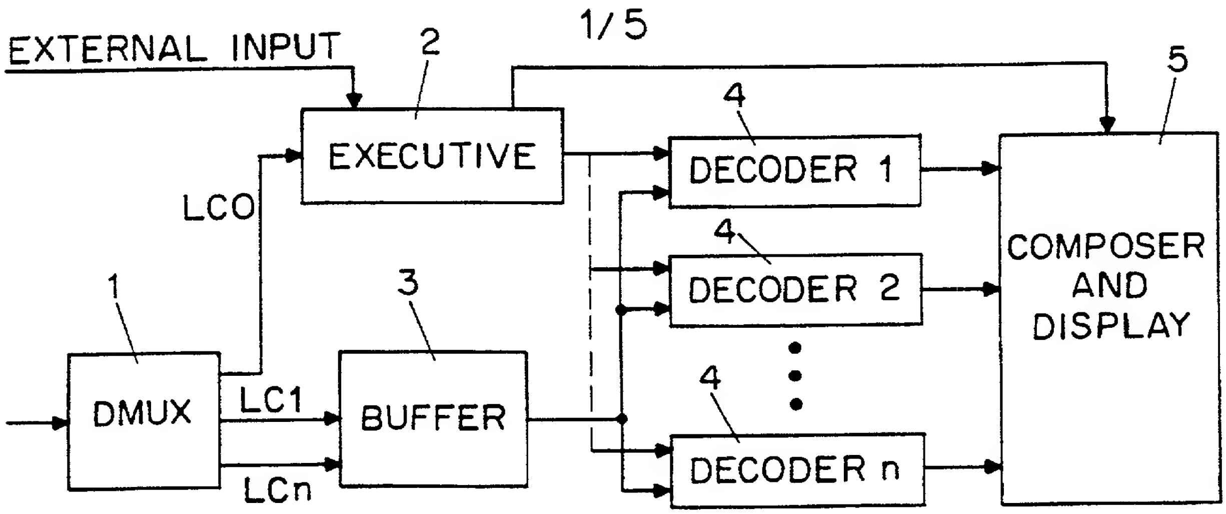

FIG. 1

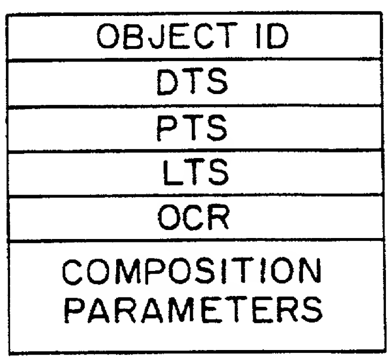

FIG. 2a

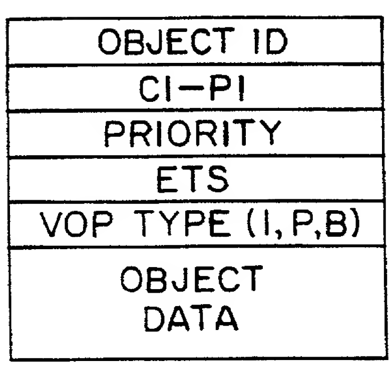

FIG. 2b

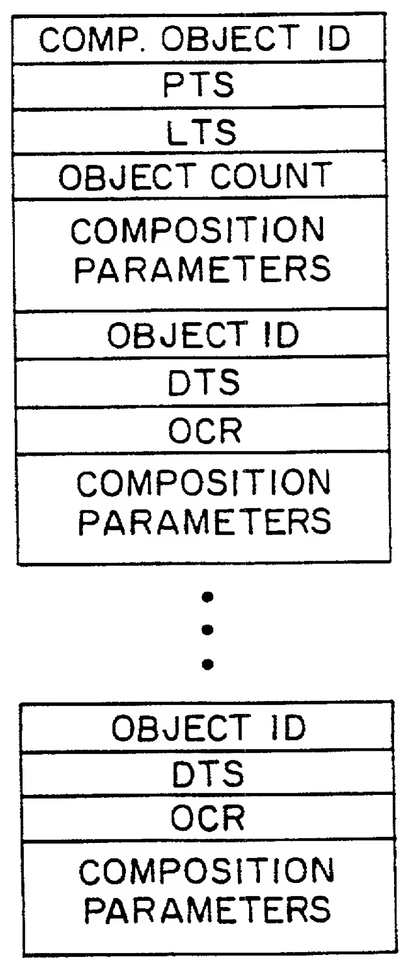

FIG. 2c

Node ID	Node param	• • •	Node param

Comp Obj ID (0X0001)	Root node (RN)	No. children	Child node (CN)	• • •	Child node (CN)

FIG. 3

Comp Obj (0X0002)	Table size	Stream ID	Node ID	• • •	Stream ID	Node ID

FIG. 4

Comp Obj (0X0003)	Update type (0x00-delete	Node ID

Comp Obj (0X0003)	Update type (0x01-add)	Parent Node ID	Node

Comp Obj (0X0003)	Update type (0x00-modify	Node ID	No. Param	Param index	Value	• • •	Param index	Value

FIG. 5

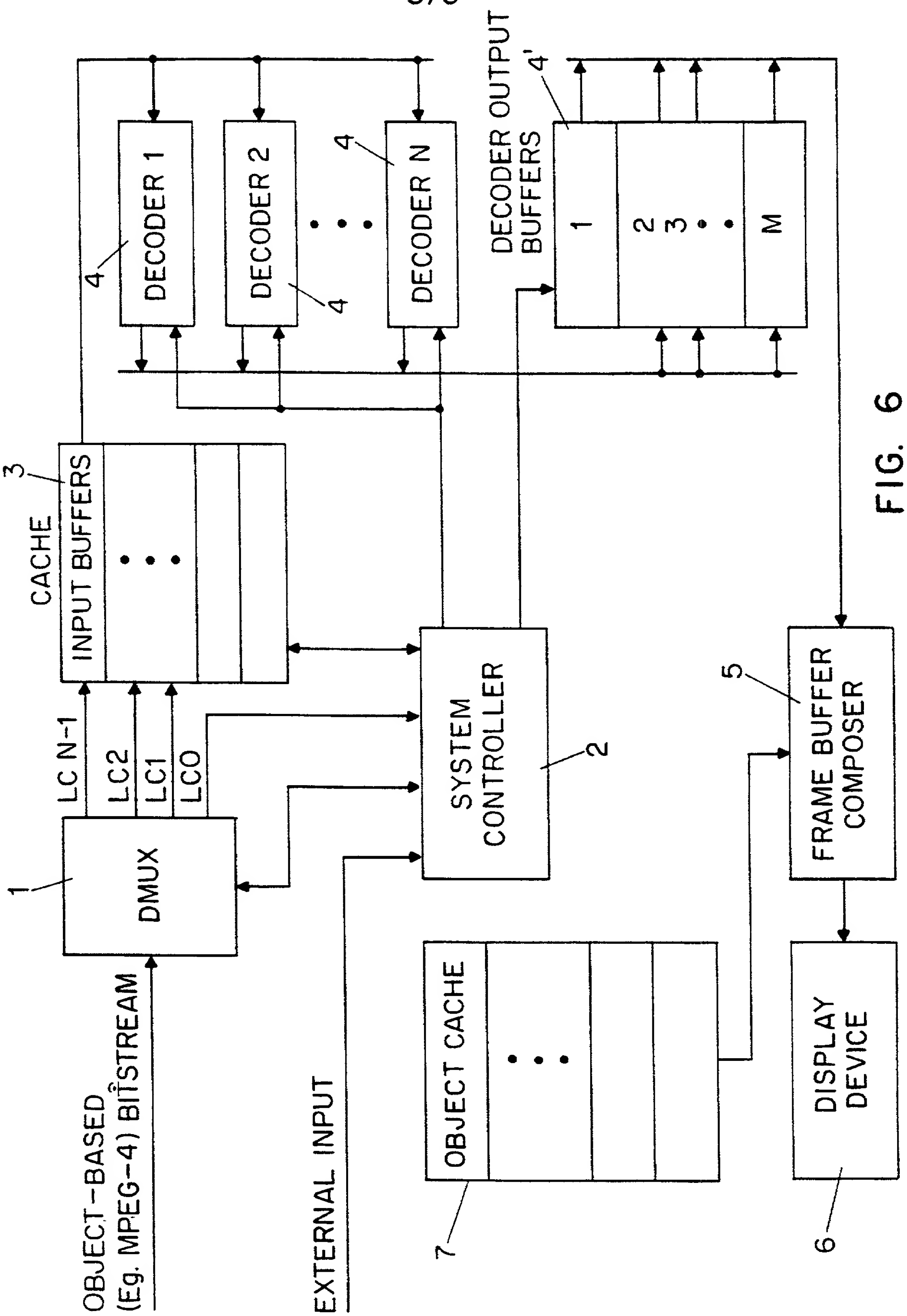
OBJECT-BASED (Eg. MPEG-4) BITSTREAM
DMUX
1
LC N-1
LC2
LC1
LC0
CACHE
INPUT BUFFERS
3
DECODER 1
DECODER 2
DECODER N
4
4'
DECODER OUTPUT BUFFERS
1
2
3
M
EXTERNAL INPUT
SYSTEM CONTROLLER
2
FRAME BUFFER COMPOSER
5
OBJECT CACHE
7
DISPLAY DEVICE
6

FIG. 6

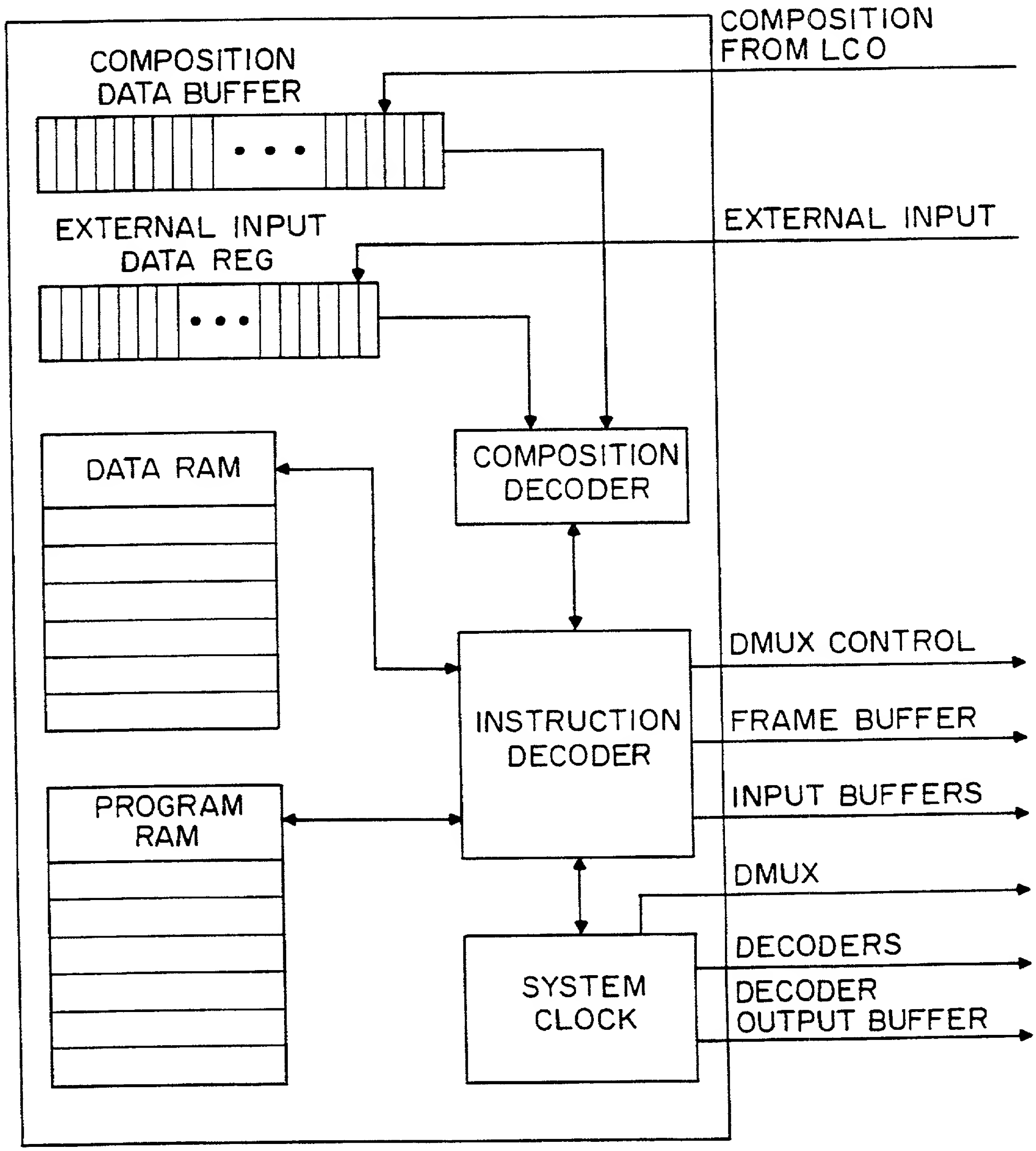
COMPOSITION FROM LCO
COMPOSITION DATA BUFFER
EXTERNAL INPUT
EXTERNAL INPUT DATA REG
DATA RAM
COMPOSITION DECODER
DMUX CONTROL
INSTRUCTION DECODER
FRAME BUFFER
INPUT BUFFERS
PROGRAM RAM
DMUX
SYSTEM CLOCK
DECODERS
DECODER OUTPUT BUFFER

FIG. 7

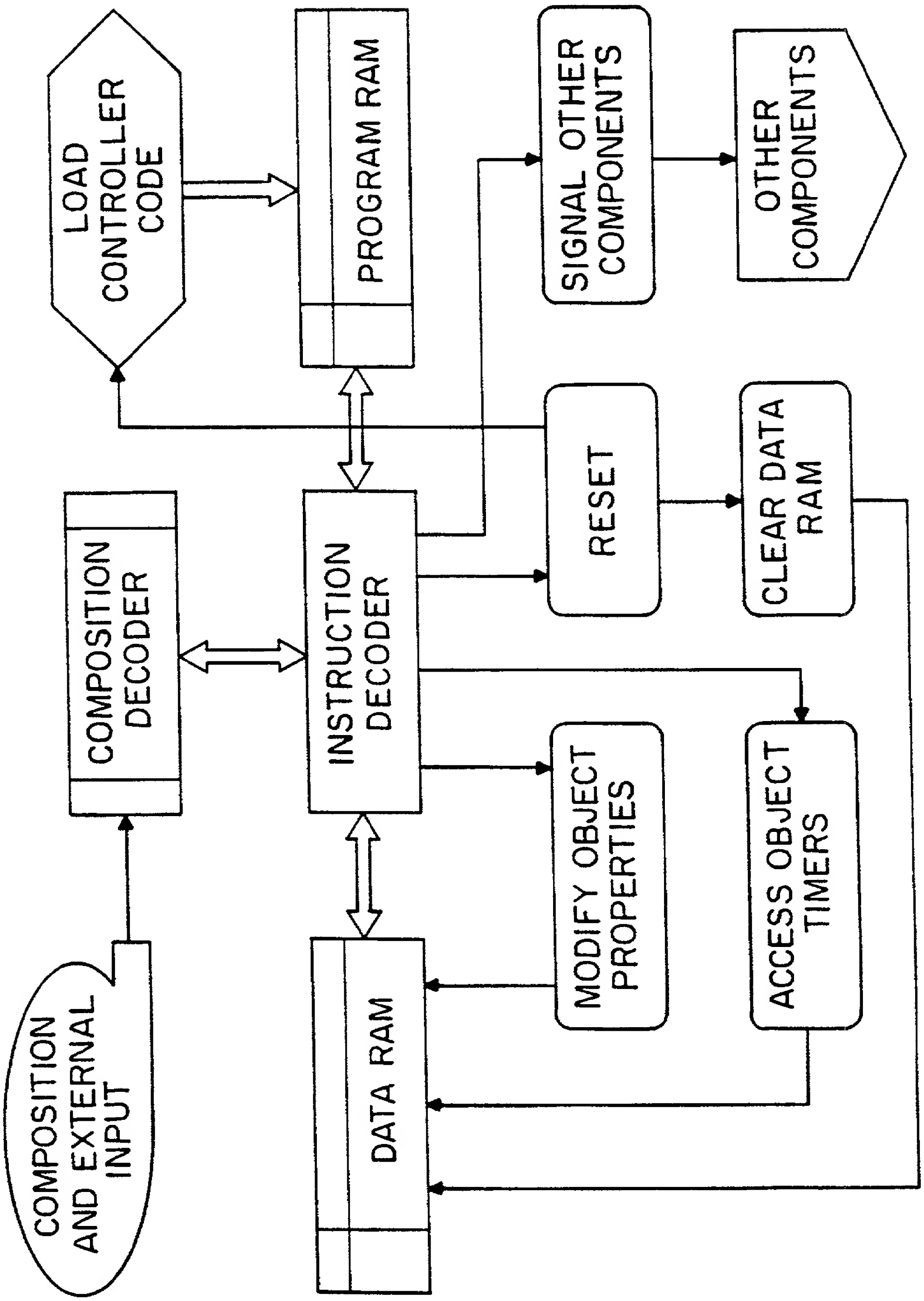
LOAD CONTROLLER CODE
PROGRAM RAM
SIGNAL OTHER COMPONENTS
OTHER COMPONENTS
COMPOSITION DECODER
INSTRUCTION DECODER
RESET
CLEAR DATA RAM
COMPOSITION AND EXTERNAL INPUT
DATA RAM
MODIFY OBJECT PROPERTIES
ACCESS OBJECT TIMERS

FIG. 8

OIPE
JAN 1 3 2000
JC67
PATENT & TRADEMARK OFFICE

BAKER & BOTTS, L.L.P.
FILE NO.: A30919 PCT USA

COMBINED DECLARATION AND POWER OF ATTORNEY

(Original, Design, National Stage of PCT, Divisional, Continuation or C-I-P Application)

As a below named inventor, I hereby declare that:
My residence, post office address and citizenship are as stated below next to my name; I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

OBJECT-BASED AUDIO-VISUAL TERMINAL AND BITSTREAM STRUCTURE

This declaration is of the following type:

- [] original
- [] design
- [X] national stage of PCT
- [] divisional
- [] continuation
- [] continuation-in-part (C-I-P)

the specification of which: *(complete (a), (b), or (c))*

(a) [ ] is attached hereto.
(b) [X] was filed on August 13, 1999 as Application Serial No. 09/367,433 and was amended on *(if applicable).*
(c) [X] was described and claimed in PCT International Application No. PCT/US98/02668 filed on February 13, 1998 and was amended on *(if applicable).*

Acknowledgement of Review of Papers and Duty of Candor

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of the subject matter claimed in this application in accordance with Title 37, Code of Federal Regulations § 1.56.

[ ] In compliance with this duty there is attached an information disclosure statement. 37 CFR 1.98.

Priority Claim

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT International Application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT International Application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application on which priority is claimed

(complete (d) or (e))

(d) [X] no such applications have been filed.
(e) [ ] such applications have been filed as follows:

PRIOR FOREIGN/PCT APPLICATION(S) FILED WITHIN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION			
COUNTRY APPLICATION NO.	DATE OF FILING (day, month, year)	DATE OF ISSUE (day, month, year)	PRIORITY CLAIMED UNDER 35 USC 119
			[ ] YES NO [ ]
			[ ] YES NO [ ]
			[ ] YES NO [ ]
ALL FOREIGN APPLICATION[S], IF ANY, FILED MORE THAN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO SAID APPLICATION			
			[ ] YES NO [ ]
			[ ] YES NO [ ]
			[ ] YES NO [ ]

Claim for Benefit of Prior U.S. Provisional Application(s)

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below:

Provisional Application Number	Filing Date
60/037,779	February 14, 1997

Claim for Benefit of Earlier U.S./PCT Application(s) under 35 U.S.C. 120

(complete this part only if this is a divisional, continuation or C-I-P application)

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) in the manner provided by the first paragraph of Title 35, United States Code § 112, I acknowledge the duty to disclose information as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

(Application Serial No.) (Filing Date) (Status) (patented, pending, abandoned)

(Application Serial No.) (Filing Date) (Status) (patented, pending, abandoned)

Power of Attorney

As a named inventor, I hereby appoint Dana M. Raymond, Reg. No. 18,540; Frederick C. Carver, Reg. No. 17,021; Francis J. Hone, Reg. No. 18,662; Joseph D. Garon, Reg. No. 20,420; Arthur S. Tenser, Reg. No. 18,839; Ronald B. Hildreth, Reg. No. 19,498; Thomas R. Nesbitt, Jr., Reg. No. 22,075; Robert Neuner, Reg. No. 24,316; Richard G. Berkley, Reg. No. 25,465; Richard S. Clark, Reg. No. 26,154; Bradley B. Geist, Reg. No. 27,551; James J. Maune, Reg. No. 26,946; John D. Murnane, Reg. No. 29,836; Henry Tang, Reg. No. 29,705; Robert C. Scheinfeld, Reg. No. 31,300; John A. Fogarty, Jr., Reg. No. 22,348; Louis S. Sorell, Reg. No. 32,439; Rochelle K. Seide Reg. No. 32,300; Gary M. Butter, Reg. No. 33,841; Marta E. Delsignore, Reg. No. 32,689; and Lisa B. Kole, Reg. No. 35,225 of the firm of BAKER & BOTTS, L.L.P., with offices at 30 Rockefeller Plaza, New York, New York 10112, as attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith

SEND CORRESPONDENCE TO:	DIRECT TELEPHONE CALLS TO:
BAKER & BOTTS, L.L.P. 30 ROCKEFELLER PLAZA, NEW YORK, N.Y. 10112 CUSTOMER NUMBER: 21003	BAKER & BOTTS, L.L.P. (212) 705-5000

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge

that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR	LAST NAME 1-00 Eleftheriadis	FIRST NAME Alexandros	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY New York NY	STATE or FOREIGN COUNTRY New York	COUNTRY OF CITIZENSHIP Greece	
POST OFFICE ADDRESS	POST OFFICE ADDRESS Dept. of Electrical Engineering, Columbia University	CITY New York	STATE or COUNTRY New York	ZIP CODE 10027
DATE 12/28	SIGNATURE OF INVENTOR			
FULL NAME OF SECOND JOINT INVENTOR, IF ANY	LAST NAME 2-00 Kalva	FIRST NAME Hari	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY New York NY	STATE or FOREIGN COUNTRY New York	COUNTRY OF CITIZENSHIP India	
POST OFFICE ADDRESS	POST OFFICE ADDRESS Dept. of Electrical Engineering, Columbia University	CITY New York	STATE or COUNTRY New York	ZIP CODE 10027
DATE Dec 28 '99	SIGNATURE OF INVENTOR			
FULL NAME OF THIRD JOINT INVENTOR, IF ANY	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE or COUNTRY	ZIP CODE
DATE	SIGNATURE OF INVENTOR			
FULL NAME OF FOURTH JOINT INVENTOR, IF ANY	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE or COUNTRY	ZIP CODE
DATE	SIGNATURE OF INVENTOR			
FULL NAME OF FIFTH JOINT INVENTOR, IF ANY	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE or COUNTRY	ZIP CODE
DATE	SIGNATURE OF INVENTOR			
FULL NAME OF SIXTH JOINT INVENTOR, IF ANY	LAST NAME	FIRST NAME	MIDDLE NAME	
RESIDENCE & CITIZENSHIP	CITY	STATE or FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP	
POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE or COUNTRY	ZIP CODE